



The semi-automated algorithm for the detection of bone marrow oedema lesions in patients with axial spondyloarthritis

Iwona Kucybała¹ · Zbysław Tabor² · Jakub Polak¹ · Andrzej Urbanik¹ · Wadim Wojciechowski¹

Received: 4 December 2019 / Accepted: 2 January 2020 / Published online: 18 January 2020
© The Author(s) 2020

Abstract

The aim of the study was to create the efficient tool for semi-automated detection of bone marrow oedema lesions in patients with axial spondyloarthritis (axSpA). MRI examinations of 22 sacroiliac joints of patients with confirmed axSpA-related sacroiliitis (median SPARCC score: 14 points) were included into the study. Design of our algorithm is based on Maksymowych et al. evaluation method and consists of the following steps: manual segmentation of bones (T1W sequence), automated detection of reference signal region, sacroiliac joint central lines and ROIs, a division of ROIs into quadrants, automated detection of inflammatory changes (STIR sequence). As a gold standard, two sets of manual lesion delineations were created. Two approaches to the performance assessment of lesion detection were considered: pixel-wise (detections compared pixel by pixel) and quadrant-wise (quadrant to quadrant). Statistical analysis was performed using Spearman's correlation coefficient. Correlation coefficient obtained for pixel-wise comparison of semi-automated and manual detections was 0.87 ($p=0.001$), while for quadrant-wise analysis was 0.83 ($p=0.001$). The correlation between two sets of manual detections was 0.91 for pixel-wise comparison ($p=0.001$) and 0.88 ($p=0.001$) for quadrant-wise approach. Spearman's correlation between two manual assessments was not statistically different from the correlation between semi-automated and manual evaluations, both for pixel- ($p=0.14$) and quadrant-wise ($p=0.17$) analysis. Average single slice processing time: 0.64 ± 0.30 s. Our method allows for objective detection of bone marrow oedema lesions in patients with axSpA. The quantification of affected pixels and quadrants has comparable reliability to manual assessment.

Keywords Spondylarthritis · Sacroiliitis · Sacroiliac joint · Ankylosing spondylitis · Magnetic resonance imaging · Algorithms · Diagnostic imaging

This study will be presented on European Congress of Radiology in Vienna, Austria (11–15th March 2020).

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00296-020-04511-w>) contains supplementary material, which is available to authorized users.

✉ Wadim Wojciechowski
wadim.wojciechowski@uj.edu.pl

¹ Department of Radiology, Jagiellonian University Medical College, 19 Kopernika Street, 31-501 Kraków, Poland

² Department of Biocybernetics and Biomedical Engineering, AGH University of Science and Technology, 30 Adama Mickiewicza Avenue, 30-059 Kraków, Poland

Introduction

Axial spondyloarthritis (axSpA) is an inflammatory rheumatic condition, involving primarily an axial skeleton and progressively leading to the sacroiliac, intervertebral and facet joint immobilization [1]. The definition of active sacroiliitis in MRI, fulfilling the ASAS (Assessment in SpondyloArthritis international Society) criteria, is bone marrow oedema visible on T2-weighted sequence sensitive for free water (e.g., STIR sequence) or bone marrow enhancement (osteitis) present on T1-weighted sequence after contrast media administration. The lesion should be located periarticularly in a subchondral bone, and must be visible on the two consecutive slices of MRI examination or on only one slice if at least two lesions are noticeable [2]. Nonetheless, the assessment of active sacroiliitis in MRI is not an easy task, especially when lesions are small. The inter-rater agreement on the presence of 'positive

MRI' according to ASAS criteria is substantial ($\kappa=0.73$), yet it is still unsatisfactory [3].

Computer-aided detection (CAD) is one of the fastest developing technologies in Radiology. Fundamental advantages of these systems in Radiology are a shortening of an examination assessment time, an increase in radiologists' productivity as well as a reduction of errors incidence due to enlarged objectivity of examination evaluation [4]. CAD has not gained much popularity in Rheumatology as yet. The majority of existing solutions focuses on the evaluation of some features of such diseases as rheumatoid arthritis (RA) and osteoarthritis (OA). Regarding axSpA, only one tool has been developed, which enables for the semi-automated quantification of active sacroiliitis in MRI [5]. However, the primary drawback of this method is that it requires manual selection of lesions (software only detects their contours) and does not detect lesions missed by the observer [5]. So far, any software allowing for the detection of bone marrow oedema lesions in MRI, which do not require their manual selection, has not been designed yet, even though the need is large [6].

The aim of our study was to create the efficient tool for the semi-automated detection of bone marrow oedema lesions in patients with axSpA.

Material and methods

Material

The study obtained an approval from the Institutional Bioethics Committee (No. of approval: 1072.6120.156.2018, date of approval: 22nd June 2018).

MRI examinations of 22 sacroiliac joints of patients with confirmed sacroiliitis in the course of axSpA were included into this retrospective study.

Methods

Imaging protocol

All examinations were performed with using 3.0 Tesla MRI scanner (Achieva, Philips Healthcare, Amsterdam, The Netherlands) and 8-channel phased-array XL-torso body matrix coil. To further analysis, only T1-weighted and STIR (short tau inversion recovery) sequences were included. Both sequences were acquired in the coronal oblique plane, parallel to the long axis of sacral bone and the position of patient remained unchanged during their acquisition time. Detailed scan parameters were:

- for T1-weighted turbo spin echo (TSE) sequence—TR 500 ms, TE 14 ms, flip angle 90, NEX 1, slice thickness 3 mm, matrix 560×560 , FOV $240 \times 240 \times 71$,
- for STIR TSE sequence—TR 5239 ms, TE 30 ms, inversion time 190 ms, flip angle 90, NEX 2, slice thickness 3 mm, matrix 400×400 , FOV $240 \times 240 \times 71$.

Sample 2D image from T1-weighted sequence (Fig. 1a), with corresponding image from STIR sequence (Fig. 1b), are presented in Fig. 1.

Image processing

Each three-dimensional (3D) MRI examination analysed in the present study consists of a sequence of various number of two-dimensional (2D) slices (between 18 and 24), which were assessed separately. The design of our algorithm was based on the method of systematic evaluation of active inflammatory changes in sacroiliac joints described by Maksymowych et al. [7].

The semi-automated algorithm for the detection of bone marrow oedema consisted of following steps:

1. Manual segmentation of the sacral bone and visible parts of both iliac bones on T1-weighted sequence images in Segmentation Editor plugin for ImageJ (National Institutes of Health, Bethesda, MD, USA). A unique label was assigned to each of bones.
2. Detection of reference signal region.

The central axis of the sacral bone was found and all sacrum pixels, which were closer to the central axis than some user-selected distance REF_{TH} were marked as belonging to the reference signal region.

3. Detection of sacroiliac joint central lines.

The distance to the iliac and sacral bone for each non-bone pixel was calculated. Then, an absolute value of the difference of these two distances was assigned to each of non-bone pixels. Pixels at the central line of the joint were equidistant to sacral and iliac bone, and therefore zero values were assigned to them. Next, Dijkstra's shortest path algorithm was used to find central joint lines.

4. Detection of regions of interest (ROIs).

ROIs were defined as parts of bone adjacent to joint surfaces on some user-specified distance, where algorithm searched for inflammatory changes. First, bony borders of joint surfaces were determined, which were simply projections of sacroiliac joints central lines on the surface of both sacral and iliac bones. Subsequently, separately for each bone, distances from pixels composing the bone to its joint surface were calculated. All pixels for which the distance was less than 10 mm were assigned to ROI of this bone.

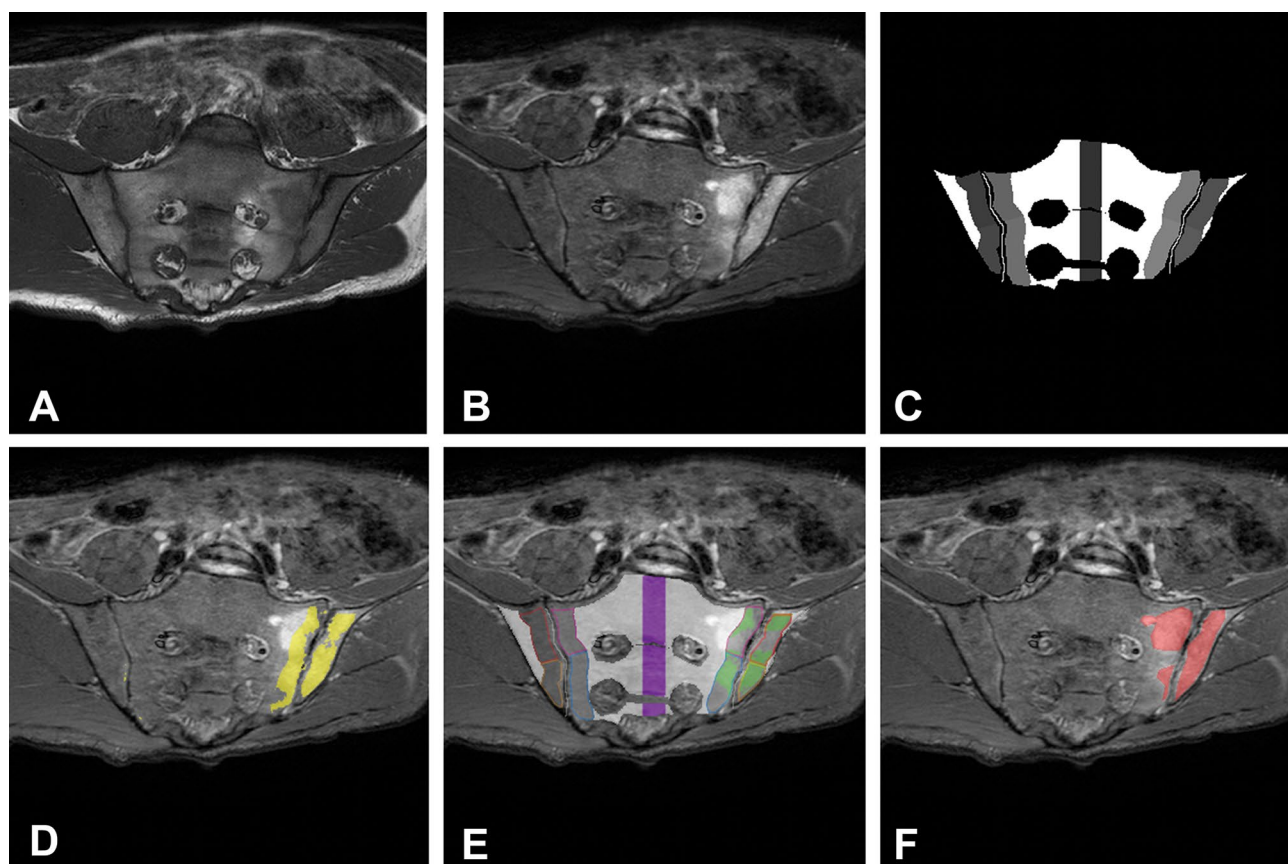


Fig. 1 **a** A sample image from T1-weighted sequence, **b** corresponding image from STIR sequence, **c** a mask containing segmentation of bones with marked reference signal region (vertical dark grey strip) and 8 quadrants (brighter grey tones), **d** results of bone marrow

oedema detection (yellow) by our algorithm for $R=1000$, $REF_{TH}=10$ and $TH_{OPT}=1.5$, **e** for $R=1000$, $REF_{TH}=10$ and $TH_{OPT}=3.5$ (green) with marked reference region (purple) and quadrants, **f** and manual, by the radiologist (red)

5. Division of ROIs into quadrants.

A midpoint of the central line of each sacroiliac joint was set and a straight line, perpendicular to the central line of the joint and passing through its midpoint, was determined to divide ROIs into upper and lower quadrants.

6. Detection of inflammatory changes.

Because the position of patient during the acquisition of T1-weighted and STIR sequence was unchanged, the reference region and quadrants determined on T1-weighted sequence images were transferred into STIR sequence images. A sample mask containing segmented bones with marked reference signal region, sacroiliac joint central lines, and quadrants is shown in Fig. 1c. To detect bone marrow oedema on STIR sequence images, a set of R reference pixels from the reference region was assigned to each pixel within ROI. Next, mean and standard deviation of the signal intensity of the reference set was computed. Subsequently, we calculated the test statistics, equal to the difference of

signal intensity in the tested pixel and mean intensity of the reference set divided by the standard deviation of the reference set. If the test statistics exceeded a user-selected threshold TH , the presence of bone marrow oedema was detected in the tested pixel.

Only the first step was performed manually, steps 2–6 were fully automated. Detailed description of image processing in steps 2–6 is presented in Online Resource 1.

Manual detection of lesions

As a gold standard for comparison of our algorithm's performance, two sets of manual delineations of bone marrow oedema lesions of all patients were created using Segmentation Editor plugin for ImageJ. All manual segmentations were produced by two independent radiologists, who did not contact with each other during the process of lesion identification.

Statistical analysis

To evaluate the results of semi-automated detection of bone marrow oedema, they were compared with the results of manual detection. Two approaches to the performance assessment of semi-automated lesion detection were considered. First, in a pixel-wise analysis, manual and semi-automated detections were compared pixel by pixel and pixel-wise false-positive and true-positive rates were determined in the function of the threshold TH. Second, in a quadrant-wise analysis, the manual and semi-automated detections were compared quadrant to quadrant. Optimal values of REF_{TH} , R , and TH for semi-automated lesion detection were determined based on ROC curve analysis and Youden's statistics. Manual detections of independent readers were also compared in pixel- and quadrant-wise manner. The normality of data distribution was evaluated prior to correlation analysis with the use of Shapiro–Wilk test. Basing on the results of the normality test, correlation analysis was conducted with the use of Spearman's rank correlation coefficient to determine the association between the results of different measurement series (semi-automated vs. manual and manual vs. manual). Statistical analysis is described in details in Online Resource 2.

Results

Overall, 54.5% ($n=6$) of patients included in the study were male. The median age of patients was 31 years (range 18–38 years). The median SPARCC score of patients was 14 points (range 2–89). Inflammatory back pain was present in 90.9% of our patients ($n=10$), peripheral arthritis in 45.5% of them ($n=5$) and HLA-B27 haplotype was detected in 27.3% ($n=3$) of them. Moreover, 18.2% ($n=2$) of patients had a family history of SpA, 18.2% ($n=2$) of them suffer from enthesitis and 9.1% ($n=1$) of them had inflammatory bowel disease.

First, we analysed ROC curves (Fig. A1 in Online Resource 3) generated for the basic version of our algorithm, designed strictly according to standard procedure proposed by Maksymowych et al.—that is, the reference signal in the midline of the sacral bone was determined at the same level and slice as analysed pixel located in ROI. Generally,

AUC was enlarging with increasing REF_{TH} and R , and the values $R=1000$ points, $REF_{TH}=10$ mm were identified as the most optimal. The area under the pixel-wise ROC curve corresponding to $REF_{TH}=10$ and $R=1000$ was equal to 0.899 [standard error (SE) < 0.001] for Reader 1 and 0.879 (SE < 0.001) for Reader 2. Next, the area under the quadrant-wise ROC curve, also for $REF_{TH}=10$ and $R=1000$, equalled to 0.876 (SE = 0.016) for Reader 1 and 0.870 (SE = 0.016) for Reader 2. There was no statistically significant difference between the area under the pixel-wise and the quadrant-wise ROC curves ($p=0.16$ for comparison of the biggest and the smallest aforementioned AUC). In the next step, Youden's statistics in function of the threshold value TH_{OPT} were calculated for pixel- and quadrant-wise ROC curves (Fig. A2 in Online Resource 3). Set values of optimal thresholds, their sensitivity, and specificity are shown in Table 1.

Further, the correlation between semi-automated and manual detections was evaluated. First, Shapiro–Wilk test was performed to assess normality and neither the number of pixels nor quadrants with inflammatory lesions had normal distribution ($p < 0.001$ in all cases). For this reason, Spearman's rank correlation coefficient was used both for pixel- and quadrant-wise analysis. The results of semi-automated bone marrow oedema detection by our algorithm for $R=1000$, $REF_{TH}=10$ and $TH_{OPT}=1.5$ (Fig. 1d), $R=1000$, $REF_{TH}=10$ and $TH_{OPT}=3.5$ (Fig. 1e) and manual, by the radiologist (Fig. 1f) are depicted on Fig. 1. Graphs showing the total number of pixels/quadrants with lesions detected by our algorithm for $R=1000$, $REF_{TH}=10$ and $TH_{OPT}=1.5$ (pixel-wise comparison) as well as $TH_{OPT}=3.5$ (for quadrant-wise comparison) plotted against the total number of pixels/quadrants with lesions detected manually are presented in Fig. 2. The correlation between semi-automated and manual detections was slightly stronger for the pixel-wise than for quadrant-wise analysis—details regarding the strength of correlation are presented in Table 2.

In several sections from MRI examination of some patients, the number of pixels in the midline of the sacral bone was not sufficient to perform the standard analysis of the reference signal and therefore the detection of inflammatory lesions on these sections could not be performed. In consequence, we decided to develop the second method of the reference signal calculation, namely, as an average

Table 1 Optimal thresholds, their sensitivity, and specificity for both methods of reference signal calculation

Reference signal range	Comparison type	Optimal threshold (TH_{OPT})	Sensitivity		Specificity	
			Reader 1 (%)	Reader 2 (%)	Reader 1 (%)	Reader 2 (%)
Limited set of reference pixels	Pixel-wise	1.5	86	79	86	89
	Quadrant-wise	3.5	80	75	86	88
Reference signal from the entire reference region	Pixel-wise	1.5	87	78	90	92
	Quadrant-wise	2.5	89	82	81	84

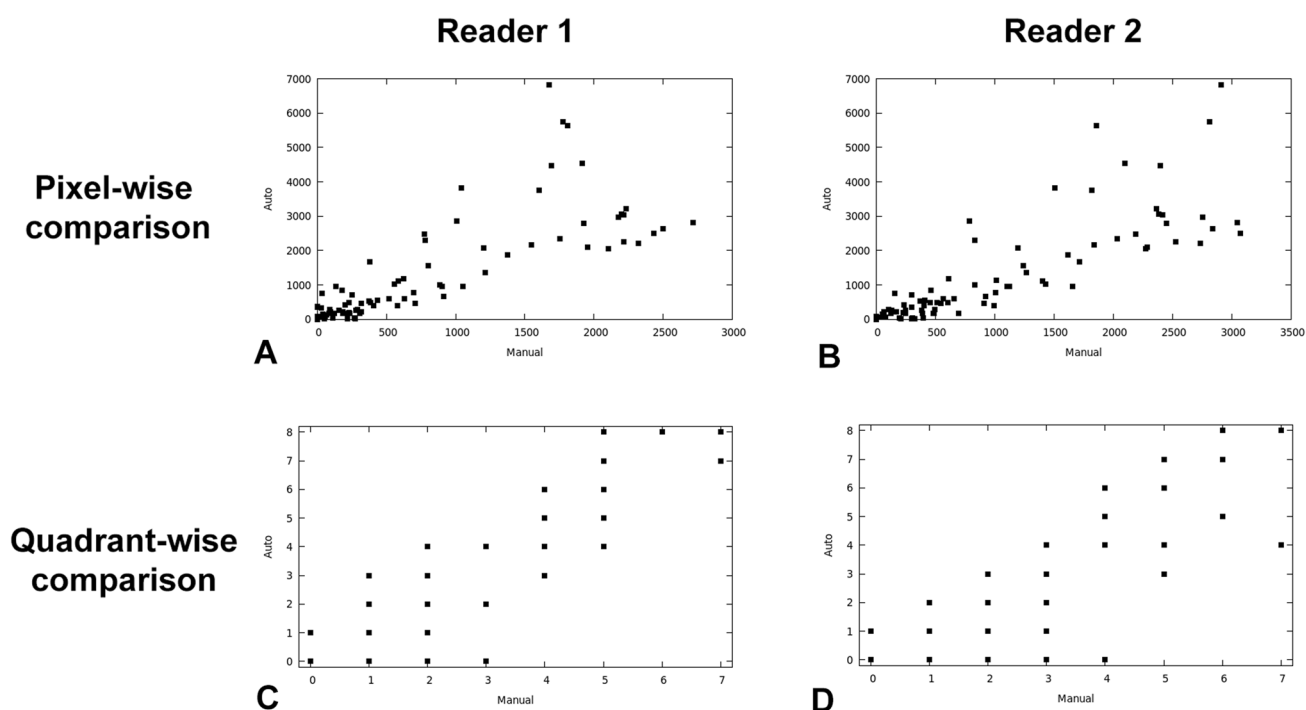


Fig. 2 **a, b** The total number of pixels with lesions as detected by the algorithm for $R=1000$, $REF_{TH}=10$ and $TH_{OPT}=1.5$, plotted against the total number of pixels with lesions as detected manually by **a** Reader 1 and **b** 2. **c, d** The total number of quadrants with lesions as

detected by the algorithm for $R=1000$, $REF_{TH}=10$ and $TH_{OPT}=3.5$, plotted against the total number of quadrants with lesions as detected manually by **c** Reader 1 and **d** 2

Table 2 Correlation between different detection methods both for pixel- and quadrant-wise comparisons

	Comparison type	<i>p</i> value	Correlation coefficient (<i>r</i>)		95% confidence interval of <i>r</i>	
			Reader 1	Reader 2	Reader 1	Reader 2
Semi-automated vs. manual (limited reference region)	Pixel-wise	0.001	0.86	0.87	0.79–0.91	0.83–0.92
	Quadrant-wise	0.001	0.83	0.82	0.75–0.89	0.73–0.88
Semi-automated vs. manual (whole reference region)	Pixel-wise	0.001	0.86	0.90	0.79–0.91	0.85–0.93
	Quadrant-wise	0.001	0.77	0.76	0.66–0.85	0.65–0.84
Manual (reader 1) vs. manual (reader 2)	Pixel-wise	0.001	0.91		0.86–0.94	
	Quadrant-wise	0.001	0.88		0.81–0.92	

of the reference signal from the entire reference region R_{REF} of some user-selected thickness REF_{TH} . For this method, ROC curves generated for $REF_{TH}=10$ mm, both for pixel- and quadrant-wise analysis, provided the biggest AUCs. The area under the pixel-wise ROC curve was equal to 0.925 ($SE<0.001$) for Reader 1 and 0.897 ($SE<0.001$) for Reader 2, while under the quadrant-wise ROC curve was at the level of 0.904 ($SE=0.014$) for Reader 1 and 0.889 ($SE=0.014$) for Reader 2. Optimal thresholds, their sensitivity, and specificity for pixel- and quadrant-wise comparison are presented in Table 1. In this case, correlation between semi-automated and manual detections was slightly higher for the pixel-wise analysis than for

quadrant-wise comparison, but as in the first method, the difference is minimal—detailed results are grouped in Table 2. This method of reference signal calculation enabled to obtain an improvement of algorithm performance for pixel-wise comparison ($p<0.001$ for AUCs difference), as opposed to the quadrant-wise comparison, for which the difference was not statistically significant ($p\geq 0.05$ for AUC difference).

The correlation between manual detections of bone marrow oedema lesions performed by two independent observers was also evaluated in pixel- and quadrant-wise manner (Fig. 3). Spearman's correlation coefficients for pixel- and quadrant-wise comparisons are presented in Table 2.

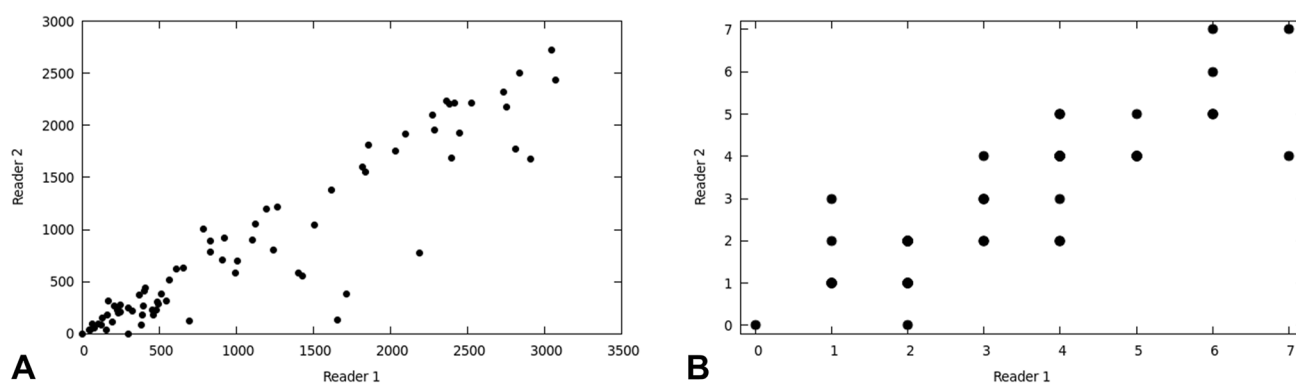


Fig. 3 The total number of pixels (a) and quadrants (b) with lesions as detected by two independent manual readers

Spearman's correlation coefficient for comparison between the two manual assessments in terms of the number of lesion pixels was not statistically different from coefficients achieved for the comparison between semi-automated and manual detections for both methods of reference signal calculation (0.91 vs. 0.86 for the first method, 0.91 vs. 0.86 for the second method; $p=0.140$ in both cases). Moreover, there was not any statistically significant difference in Spearman's coefficient between the one determined for comparison of two manual assessments of the number of affected quadrants and achieved for comparison of semi-automated and manual assessments for the basic method of reference signal calculation (0.88 vs. 0.82, $p=0.170$). In the case of the second method of reference signal computation (over the entire reference region), Spearman's correlation coefficient for quadrant-wise comparison of two manual delineations was significantly higher than for the comparison of manual and semi-automated detections (0.88 vs. 0.76, $p=0.020$).

The average processing time for a single slice was equal to 0.64 ± 0.30 s for $R=1000$ and below 0.1 s for an entire R_{REF} set for $\text{REF}_{\text{TH}}=10$ mm on a PC with 8 GB RAM and Intel® Core i7-5500U 2.40 GHz CPU, without including the time required for manual segmentation of iliac and sacral bones (30–45 min).

Discussion

The automation of the process of bone marrow oedema detection is a highly difficult task. The definition of the bone marrow oedema lesion developing in the course of axSpA [2] seems to be specific, yet after thorough analysis it is obvious that there are some issues that remain unclear, such as the exact interpretation of the term 'highly suggestive of axSpA'. To increase reliability of sacroiliitis assessment, Maksymowych et al. created SPARCC (SpondyloArthritis Research Consortium of Canada) score and disambiguated the rules of active sacroiliitis evaluation [7]. Nonetheless, it

is undoubtedly difficult to perform the examination assessment with the use of SPARCC scoring method in a reasonable time and completely objectively using only the human eye. However, these objectives were met by our semi-automated algorithm, which credibility was proven to be comparable to the results of manual assessments. The biggest similarity to the agreement between two manual readers was observed for the detection of inflammatory lesion pixels regardless of the method of reference signal acquisition as well as for the detection of quadrants affected by bone marrow oedema using the basic method of reference signal calculation. Both detection of all pixels belonging to the potential lesion and quadrants affected by the inflammation have similar performance, but their application could be different. Focusing detection on particular pixels enables to detect and highlight areas suspected of bone marrow oedema presence and therefore radiologist can verify the actual significance of detected changes and easily describe examination results. Concentration on the detection of altered quadrants may be a good tool for a quick quantitative assessment of inflammation extent as well as a valuable help to speed up SPARCC score calculation. Both SPARCC score and the number of affected quadrants could be used as a quantitative measure to monitor disease activity during the treatment or clinical studies, as both these techniques have comparable reliability (ICC for the number of affected quadrants: 0.47, ICC for SPARCC: 0.55) [8, 9]. Thus, these scoring systems could be used interchangeably and the most convenient method should be chosen individually.

Our analysis confirms that both methods of reference signal acquisition have similar credibility—the calculation of the mean intensity of limited number of pixels on the same level as particular pixel within ROI, as well as averaged intensity over the whole reference region. On the contrary, intensity comparison during manual assessment is highly estimated, which gives the opportunity to make mistakes in evaluation of the presence and the extent of bone marrow oedema lesions. There is a vast range of situations, which

may mislead the radiologist—for instance, the signal intensity of bone marrow in the midline of the sacral bone could be highly heterogeneous (Fig. 4a), or the lesion could consist of two parts—well-margined of very high intensity and surrounding it large, blurred, less intense area (Fig. 4b). In the first case, the radiologist could use tools provided by a medical image viewer allowing for the calculation of the mean intensity inside the user-selected ROI and then manually compare values between the suspected lesion and the midline of the sacral bone. Regrettably, this method is time-consuming, especially when a patient has many suspicious areas within the sacroiliac joints, as well as there still could be a problem of the absent midline part of the sacrum, due to its shape. In the second case, the risk of the diagnostic mistake is high, as only well-margined, highly hyperintense part of a change might draw the attention of the evaluating radiologist, and hence the magnitude of the lesion would be underestimated. The second problem was solved by Zarco et al. tool which was invented for the semi-automated detection of bone marrow oedema borders and further lesions scoring [5]. The concept of this software is completely different from ours, as it requires the mouse-click within the suspected change to detect its actual borders, basing on pre-determined tolerance range. Hence, this software does not provide the objectivity regarding lesions detection, as an observer might unintentionally omit some small changes.

The mean analysis time per single slice was 0.64 s for reference region limited to 1000 pixels and <0.1 s for entire reference region. MRI of sacroiliac joints performed in our centre consists of 18–22 slices (for slice thickness equal to

3 mm) and the joint space is visible on approximately 8–12 slices. In consequence, an analysis time of the whole MRI examination of one patient is up to 10 s for our technique, which is faster than median analysis time of Zarco et al. method—28 s [5]. Processing time reduction for the second method of reference signal calculation is a result of an identical average reference signal intensity and its standard deviation for all tested pixels. Therefore, an analysis time of whole SIJ examination for this method of reference signal acquisition is up to 2 s, which is an excellent result. Nevertheless, the biggest limitation of our method is that before the automated detection of inflammatory changes it requires manual preparation of segmentations of bones forming the sacroiliac joints. This procedure is highly arduous, and it took our experts approximately between 30 and 45 min to perform these segmentations. At this point in time, it hinders the implementation of our method in daily clinical practice, but our team is currently working on an algorithm for automated detection of bones forming these joints. Up to now, any study describing segmentation method of bones in the sacroiliac joints region on MRI images has not been published. The first obstacle to the development of this technique is the fact that STIR sequence cannot be used to bone segmentation purposes, as clear boundaries are not visible between bones and some soft tissue structures, for instance the insertion of piriformis muscle to the sacrum. Preferred MRI sequence to the assessment of joints anatomical structure is T1-weighted sequence and it was used in previous studies to the automated segmentation of wrist bones [10, 11]. The simplest solution of this problem may

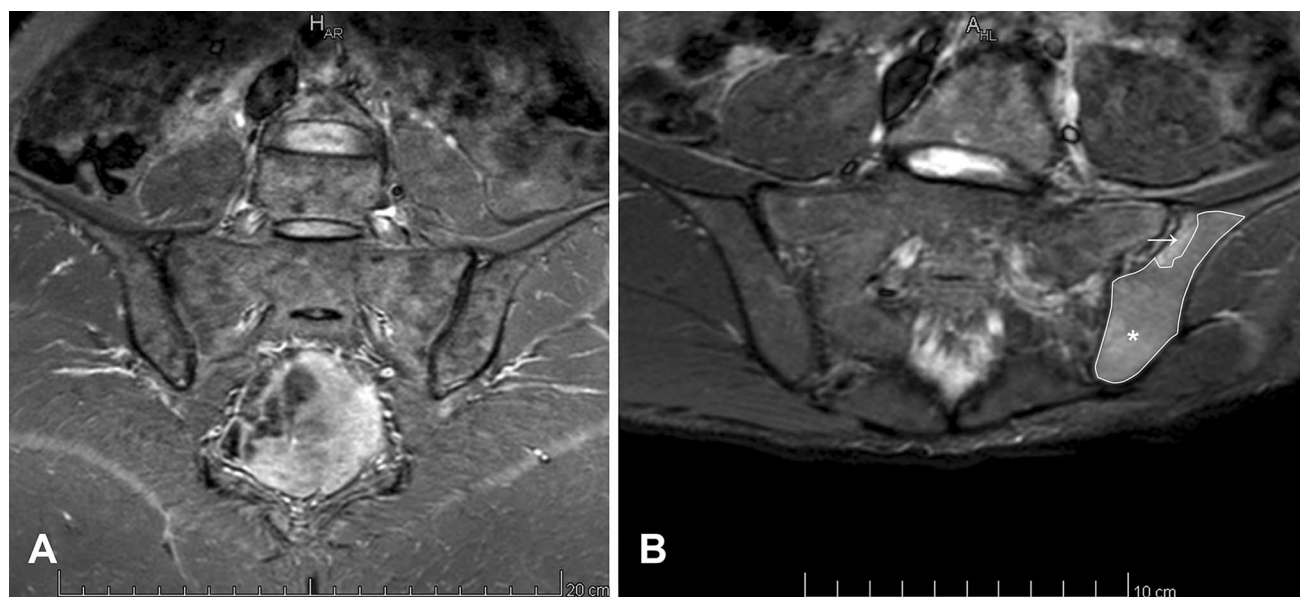


Fig. 4 **a** Heterogeneous signal intensity of bone marrow in the midline of sacral bone visible on STIR sequence. **b** Large bone marrow oedema lesion within the left iliac bone, which consists of two parts:

small, well-margined of very high intensity (arrow) and surrounding it large, blurred, less intense area (asterisk)—STIR sequence

be the detection of active sacroiliitis signs on T1-weighted sequence after contrast media administration, which is also included in ASAS criteria regarding the positive MRI definition [2]. Nonetheless, the use of gadolinium-based contrast media is linked with potential adverse events [12], and therefore sequences with contrast enhancement should be avoided when equally reliable alternative, such as STIR sequence, is available. Next obstacle is that to transfer segmentations between T1-weighted sequence and STIR, some technical parameters (as slice thickness) and patient position must be identical, what not always is achievable. Conversely, the accurate method of sacroiliac region bones segmentation and sacroiliitis detection was developed for computed tomography by Shenkman et al.—however, bones are clearly visible, hyperdense structures in this type of examination and their segmentation could be performed easily using thresholding method [13]. Moreover, computed tomography enables to visualize only late, irreversible changes within the sacroiliac joints (erosions, sclerosis, ankylosis) and is not recommended by ASAS to the diagnostics of axSpA in the early stage [14].

Currently, ASAS axSpA classification criteria consider bone marrow oedema as the only lesion which fulfils the definition of positive MRI [2] and this was the reason why we decided to design an algorithm detecting this kind of active inflammatory changes. However, next limitation of our method is that the presence of bone marrow oedema within the sacroiliac joints is not pathognomonic for axSpA and it could be visible in patients with non-specific back pain (up to 23% of cases), women in the postpartum period (21–41%), athletes (30–41%), soldiers (36%) and healthy volunteers (up to 7%) [15–19]. For this reason, investigators are searching for the perfect combination of changes within the sacroiliac joints, which will increase the specificity of axSpA diagnosis with the use of MRI. Recent reports suggest the combination of the presence of bone marrow oedema and chronic structural lesions (such as erosion, fatty infiltration or sclerosis) in at least two to three quadrants increases the specificity of axSpA diagnosis without the decline in sensitivity, in comparison to the bone marrow oedema presence alone [20, 21]. Hence, the tool for automated detection of structural changes within the sacroiliac joint should also be developed in the future.

Another limitation of our method, which is universal for all CAD systems, is the presence of the automation bias. This phenomenon is described as the tendency of over-reliance on automated systems, which leads to increased incidence of errors in examination assessment. It is the result of various factors such as the user cognitive style, previous experience with CAD systems, task complexity as well as workload, and could be avoided by strengthening user accountability [22]. Thus, it is crucial to critically review all changes highlighted by the algorithm, as well as screen the

remaining part of sacroiliac joint for the presence of inflammatory lesions omitted by the algorithm, before the decision regarding the disease is made. Finally, the last shortcoming of the study is a small number of analysed cases.

In conclusion, our semi-automated algorithm allows for highly objective and credible detection of bone marrow oedema lesions visible on MRI examination of patients with axSpA. The detection of affected pixels and quadrants with the use of our basic method has comparable reliability to manual assessment. However, further work on the algorithm is vital to automate the process of bone segmentation in the sacroiliac region.

Author contributions Conceptualization: IK, AU, WW; data collection: IK, JP; algorithm design: IK, ZT; statistical analysis: IK, ZT; writing-original draft preparation: IK; writing-review and editing: ZT, JP, AU, WW. All authors read and approved the final manuscript.

Funding None.

Compliance with ethical standards

Conflict of interest Iwona Kucybała, Zbysław Tabor, Jakub Polak, Andrzej Urbanik and Wadim Wojciechowski declare that they have no conflict of interest.

Human and animal rights All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional research committee (Jagiellonian University Bioethics Committee, No. of approval: 1072.6120.156.2018, date of approval: 22nd June 2018) and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. This article does not contain any studies with animals performed by any of the authors.

Informed consent For this type of study, informed consent is not required, because of its retrospective and non-interventional character and the use of anonymised data.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Maksymowych WP (2019) Biomarkers for diagnosis of axial spondyloarthritis, disease activity, prognosis, and prediction of response to therapy. *Front Immunol* 10:305. <https://doi.org/10.3389/fimmu.2019.00305>

2. Lambert RGW, Bakker PAC, van der Heijde D et al (2016) Defining active sacroiliitis on MRI for classification of axial spondyloarthritis: update by the ASAS MRI working group. *Ann Rheum Dis* 75:1958–1963. <https://doi.org/10.1136/annrheumdis-2015-208642>
3. van den Berg R, Lenczner G, Thévenin F, Claudepierre P, Feydy A, Reijnierse M, Saraux A, Rahmouni A, Dougados M, van der Heijde D (2015) Classification of axial SpA based on positive imaging (radiographs and/or MRI of the sacroiliac joints) by local rheumatologists or radiologists versus central trained readers in the DESIR cohort. *Ann Rheum Dis* 74:2016–2021. <https://doi.org/10.1136/annrheumdis-2014-205432>
4. Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts HJWL (2018) Artificial intelligence in radiology. *Nat Rev Cancer* 18:500–510. <https://doi.org/10.1038/s41568-018-0016-5>
5. Zarco P, Almodóvar R, Bueno Á, Molinero LM, SCAISS Study Group (2018) Development and validation of SCAISS, a tool for semi-automated quantification of sacroiliitis by magnetic resonance in spondyloarthritis. *Rheumatol Int* 38:1919–1926. <https://doi.org/10.1007/s00296-018-4104-3>
6. Lambert RG, Maksymowych WP (2019) Drs. Lambert and Maksymowych reply. *J Rheumatol* 46:542. <https://doi.org/10.3899/jrheum.181215>
7. Maksymowych WP, Dhillon SS, Chiowchanwisawakit P, Pedersen SJ, Martinez B, Østergaard M, Lambert RGW (2009) Development and validation of web-based training modules for systematic evaluation of active inflammatory lesions in the spine and sacroiliac joints in spondyloarthritis. *J Rheumatol Suppl* 84:48–57. <https://doi.org/10.3899/jrheum.090620>
8. Kucybała I, Ciuk S, Urbanik A, Wojciechowski W (2019) The usefulness of diffusion-weighted imaging (DWI) and dynamic contrast-enhanced (DCE) sequences visual assessment in the early diagnosis of axial spondyloarthritis. *Rheumatol Int* 39:1559–1565. <https://doi.org/10.1007/s00296-019-04373-x>
9. Landewé RBM, Hermann K-GA, van der Heijde DMFM, Baraliakos X, Jurik AG, Lambert RG, Østergaard M, Rudwaleit M, Salonen DC, Braun J (2005) Scoring sacroiliac joints by magnetic resonance imaging. A multiple-reader reliability experiment. *J Rheumatol* 32:2050–2055
10. Aizenberg E, Roex EAH, Nieuwenhuis WP, Mangnus L, van der Helm-van Mil AHM, Reijnierse M, Bloem JL, Lelieveldt BPF, Stool BC (2018) Automatic quantification of bone marrow edema on MRI of the wrist in patients with early arthritis: A feasibility study. *Magn Reson Med* 79:1127–1134. <https://doi.org/10.1002/mrm.26712>
11. Włodarczyk J, Wojciechowski W, Czaplicka K, Urbanik A, Tabor Z (2015) Fast automated segmentation of wrist bones in magnetic resonance images. *Comput Biol Med* 65:44–53. <https://doi.org/10.1016/j.compbiomed.2015.07.007>
12. Malikova H, Holesa M (2017) Gadolinium contrast agents - are they really safe? *J Vasc Access* 18:S1–S7. <https://doi.org/10.5301/jva.5000713>
13. Shenkman Y, Qutteineh B, Jaskowicz L, Szeskin A, Yusef A, Mayer A, Eshed I (2019) Automatic detection and diagnosis of sacroiliitis in CT scans as incidental findings. *Med Image Anal* 57:165–175. <https://doi.org/10.1016/j.media.2019.07.007>
14. Kucybała I, Urbanik A, Wojciechowski W (2018) Radiologic approach to axial spondyloarthritis: where are we now and where are we heading? *Rheumatol Int* 38:1753–1762. <https://doi.org/10.1007/s00296-018-4130-1>
15. Seven S, Østergaard M, Morsel-Carlsen L, Sørensen IJ, Bonde B, Thamsborg G, Lykkegaard JJ, Hendricks O, Jørgensen NR, Pedersen SJ (2019) MRI lesions in the sacroiliac joints for differentiation of patients with axial spondyloarthritis from postpartum women, patients with disc herniation, cleaning staff, long distance runners and healthy persons—a prospective cross-sectional study of 204 participants. *Arthritis Rheumatol* 71:2034–2046. <https://doi.org/10.1002/art.41037>
16. Varkas G, de Hooge M, Renson T et al (2018) Effect of mechanical stress on magnetic resonance imaging of the sacroiliac joints: assessment of military recruits by magnetic resonance imaging study. *Rheumatology* 57:508–513. <https://doi.org/10.1093/rheumatology/kex491>
17. Weber U, Lambert RGW, Østergaard M, Hodler J, Pedersen SJ, Maksymowych WP (2010) The diagnostic utility of magnetic resonance imaging in spondylarthritis: An international multicenter evaluation of one hundred eighty-seven subjects. *Arthritis Rheum* 62:3048–3058. <https://doi.org/10.1002/art.27571>
18. Weber U, Jurik AG, Zejden A, Larsen E, Jørgensen SH, Rufibach K, Schioldan C, Schmidt-Olsen S (2018) Frequency and anatomic distribution of magnetic resonance imaging features in the sacroiliac joints of young athletes. *Arthritis Rheumatol* 70:736–745. <https://doi.org/10.1002/art.40429>
19. Lukas C, Cyteval C, Dougados M, Weber U (2018) MRI for diagnosis of axial spondyloarthritis: major advance with critical limitations “Not everything that glitters is gold (standard)”. *RMD open* 4:e000586. <https://doi.org/10.1136/rmdopen-2017-000586>
20. Maksymowych WP, Lambert RG, Baraliakos X et al (2019) THU0361 Diagnostic performance of MRI lesions in the sacroiliac joints according to updated ASAS lesion definitions: a central reader assessment of MRI scans from the assessments in spondyloarthritis classification cohort. *Ann Rheum Dis* 78:462–463. <https://doi.org/10.1136/annrheumdis-2019-eular.6455>
21. Baraliakos X, Ghadir A, Fruth M, Kiltz U, Braun J (2019) OP0344 Which magnetic resonance imaging lesions of the sacroiliac joints are of diagnostic value for axial spondyloarthritis? *Ann Rheum Dis* 78:255–256. <https://doi.org/10.1136/annrheumdis-2019-eular.5027>
22. Goddard K, Roudsari A, Wyatt JC (2012) Automation bias: a systematic review of frequency, effect mediators, and mitigators. *J Am Med Inform Assoc* 19:121–127. <https://doi.org/10.1136/amiatnl-2011-000089>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.